JUL 1 3 2000 E

SEQUENCE LISTING

Finlay, Brett B. Kenny, Brendant Devinney, Rebekah Stein, Marcus

<120> HOST RECEPTOR FOR PATHOGENIC BACTERIA

<130> 482112.402

<140> 09/189,415

<141> 1998-11-10

<150> 60/065,130

<151> 1997-11-12

<160> 9

<170> PatentIn Ver. 2.1

<210> 1

<211> 1920

<212> DNA

<213> Escherichia coli

<400> 1

cggctgcata ccattacgtc atagtaatat aaaggaacgt gtcaaatttc taaataaaag 60 gatatatgta tgcctattgg taaccttggt aataatgtaa atggcaatca tttaattccc 120 cctgcgccgc cactaccttc acaaacagac ggcgcggcac ggggaggaac tggtcatcta 180 attagctcta caggagcatt aggatctcgt tcattgtttt ctcccctgag aaattctatg 240 gctgattctg tcgattccag agatattcca ggacttccta caaacccatc gaggcttgct 300 gcagctacat ctgagacatg cttgcttgga ggatttgaag ttctccatga taaggggcca 360 cttgatattc tcaatacgca aattggaccc tctgcatttc gtgttgaagt gcaggcagat 420 ggtactcatg ccgctattgg agaaaaaaat ggtttggagg ttagcgttac attaagtcct 480 caagaatgga gcagcttgca atctattgat actgagggta aaaacagatt tgtttttacc 540 gggggacgtg gcggtagtgg gcatccgatg gtcactgtcg catcagatat cgcggaagct 600 cgtacgaaaa tactggccaa attagaccca gacaatcatg gaggacgtca acccaaggac 660 gttgatacgc gttctgttgg tgttggcagc gcttcgggaa tagatgatgg cgttqttagc 720 gaaacccata cttcaacaac aaattccagc gttcgctcag atcctaaatt ctgggtttct 780 gtcggcgcaa ttgctgctgg tttagcggga ctggcgcaa ctggtattgc acaggcgttg 840 getttgacae eggaacegga tgateetaea aceaeegate etgateagge egeaaatget 900 gcagaaagtg caacaaaaga tcagttaacg caagaagcat tcaagaaccc tgagaaccag 960 aaagttaaca tegatgegaa eggaaatget atteegtetg gggaattaaa agatgatatt 1020 gttgagcaaa tagcacaaca agctaaagag gctggtgagg tggccagaca gcaggctgtt 1080 gaaagcaatg cacaggcgca gcagcgatat gaggatcagc atgccagacg tcaggaggaa 1140 ttacagettt categggtat tggttaegge etcageagtg cattgattgt tgetggggga 1200 attggtgctg gtgtaacgac tgcgctccat agacgaaatc agccggcaga acagacaact 1260 actacaacaa cacatacggt agtgcagcaa cagaccggag ggatacccca gcacaaggtg 1320 gcactgatgc cacaagagcg aagacgcttc tctgatagac gtgattcgca ggggagtgtt 1380 gcatcgacac actggtcaga ttcctctagc gaagtggtta atccatatgc tgaagttggg 1440 ggggctcgga atagtctatc ggctcatcag ccagaagagc atatttatga tgaggtcgct 1500 gcagatcctg gttatagcgt tattcagaat ttttcaggga gcggcccagt taccggaagg 1560 ttaataggaa ctccagggca aggattccaa agtacttatg cgcttctggc aaacagcggc 1620 ggattgcgtt taggtatggg aggattaaccg agtggtggc agacggcagt aagttctgta 1680 aatgccgcac caacgcaggg accagtacgt ttcgttaaa tatatctgtg agtatttagt 1740 tgaggttggg gtgggtggg ggggcgtttt actagcgtta atgttcaga gaacaacgtt 1800 gcagcatgg taactcttga acttctgtta ttataatcaa ttaagagaaa ttataatgtc 1860 atcaagatat gaacttttat tagataggtt tgcggaaaaa attggtgttg gatctattc 1920

<210> 2

<211> 549

<212> PRT

<213> Escherichia coli

<220>

<221> VARIANT

<222> (314)

<223> Xaa = any amino acid

<400> 2

Met Pro Ile Gly Asn Leu Gly Asn Asn Val Asn Gly Asn His Leu Ile
1 5 10 15

Pro Pro Ala Pro Pro Leu Pro Ser Gln Thr Asp Gly Ala Ala Arg Gly
20 25 30

Gly Thr Gly His Leu Ile Ser Ser Thr Gly Ala Leu Gly Ser Arg Ser 35 40 45

Leu Phe Ser Pro Leu Arg Asn Ser Met Ala Asp Ser Val Asp Ser Arg 50 55 60

Asp Ile Pro Gly Leu Pro Thr Asn Pro Ser Arg Leu Ala Ala Ala Thr 65 70 75 80

Ser Glu Thr Cys Leu Leu Gly Gly Phe Glu Val Leu His Asp Lys Gly 85 90 95

Pro Leu Asp Ile Leu Asn Thr Gln Ile Gly Pro Ser Ala Phe Arg Val 100 105 110

Glu Val Gln Ala Asp Gly Thr His Ala Ala Ile Gly Glu Lys Asn Gly
115 120 125

Leu	Glu 130	Val	Ser	Val	Thr	Leu 135	Ser	Pro	Gln	Glu	Trp 140	Ser	Ser	Leu	Gln
Ser 145	Ile	Asp	Thr	Glu	Gly 150	Lys	Asn	Arg	Phe	Val 155	Phe	Thr	Gly	Gly	Arg 160
Gly	Gly	Ser	Gly	His 165	Pro	Met	Val	Thr	Val 170	Ala	Ser	Asp	Ile	Ala 175	Glu
Ala	Arg	Thr	Arg 180	Ile	Leu	Ala	Lys	Leu 185	Asp	Pro	Asp	Asn	His 190	Gly	Gly
Arg	Gln	Pro 195	Lys	Asp	Val	Asp	Thr 200	Arg	Ser	Val	Gly	Val 205	Gly	Ser	Ala
Ser	Gly 210	Ile	Asp	Asp	Gly	Val 215	Val	Ser	Glu	Thr	His 220	Thr	Ser	Thr	Thr
Asn 225	Ser	Ser	Val	Arg	Ser 230	Asp	Pro	Lys	Phe	Trp 235	Val	Ser	Val	Gly	Ala 240
Ile	Ala	Ala	Gly	Leu 245	Ala	Gly	Leu	Ala	Ala 250	Thr	Gly	Ile	Ala	Gln 255	Ala
Leu	Ala	Leu	Thr 260	Pro	Glu	Pro	Asp	Asp 265	Pro	Thr	Thr	Thr	Asp 270	Pro	Asp
Gln	Ala	Ala 275	Asn	Ala	Ala	Glu	Ser 280	Ala	Thr	Lys	Asp	Gln 285	Leu	Thr	Gln
Glu	Ala 290	Phe	Lys	Asn	Pro	Glu 295	Asn	Gln	Lys	Val	Asn 300	Ile	Asp	Ala	Asn
Gly 305	Asn	Ala	Ile	Pro	Ser 310	Gly	Glu	Leu	Xaa	Asp 315	Asp	Ile	Val	Glu	Gln 320
Ile	Ala	Gln	Gln	Ala 325	Lys	Glu	Ala	Gly	Glu 330	Val	Ala	Arg	Gln	Gln 335	Ala
Val	Glu	Ser	Asn 340	Ala	Gln	Ala	Gln	Gln 345	Arg	Tyr	Glu	Asp	Gln 350	His	Ala
Arg	Arg	Gln 355	Glu	Glu	Leu	Gln	Leu 360	Ser	Ser	Gly	Ile	Gly 365	Tyr	Gly	Leu
Ser	Ser 370	Ala	Leu	Ile	Val	Ala 375	Gly	Gly	Ile	Gly	Ala 380	Gly	Val	Thr	Thr

Ala Leu His Arg Arg Asn Gln Pro Ala Glu Gln Thr Thr Thr Thr 385 390 395 400

Thr His Thr Val Val Gln Gln Gln Thr Gly Gly Ile Pro Gln His Lys
405 410 415

Val Ala Leu Met Pro Gln Glu Arg Arg Arg Phe Ser Asp Arg Arg Asp 420 425 430

Ser Gln Gly Ser Val Ala Ser Thr His Trp Ser Asp Ser Ser Glu 435 440 445

Val Val Asn Pro Tyr Ala Glu Val Gly Gly Ala Arg Asn Ser Leu Ser 450 455 460

Ala His Gln Pro Glu Glu His Ile Tyr Asp Glu Val Ala Ala Asp Pro 465 470 475 480

Gly Tyr Ser Val Ile Gln Asn Phe Ser Gly Ser Gly Pro Val Thr Gly 485 490 495

Arg Leu Ile Gly Thr Pro Gly Gln Gly Ile Gln Ser Thr Tyr Ala Leu 500 505 510

Leu Ala Asn Ser Gly Gly Leu Arg Leu Gly Met Gly Gly Leu Thr Ser 515 520 525

Gly Glu Thr Ala Val Ser Ser Val Asn Ala Ala Pro Thr Pro Gly 530 540

Pro Val Arg Phe Val 545

<210> 3

<211> 1723

<212> DNA

<213> Escherichia coli

<400> 3

atgcctattg gtaaccttgg tcataatccc aatgtgaata attcaattcc tcctgcacct 60 ccattacctt cacaaaccga cggtgcaggg gggcgtggtc agctcattaa ctctacgggg 120 ccgttgggat ctcgtgcgct atttacgcct gtaaggaatt ctatggctga ttctggcgac 180 aatcgtgcca gtgatgttcc tggacttcct gtaaatccga tgcgcctggc ggcgtctgag 240 ataacactga atgatggatt tgaagttctt catgatcatg gtccgctcga tactcttaac 300 aggcagattg gctcttcggt atttcgagtt gaaactcagg aagatggtaa acatattgct 360 gtcggtcaga ggaatggtgt tgagacctct gttgtttaa gtgatcaaga gtacgctcgc 420 ttgcagtcca ttgatcctga aggtaaagac aaatttgtat ttactggagg ccgtggtggt 480

gctgggcatg ctatggtcac cgttgcttca gatatcacgg aagcccqcca aaggatactg 540 gagetgttag ageccaaagg gaeeggggag tecaaaggtg etggggagte aaaaggegtt 600 ggggagttga gggagtcaaa tagcggtgcg gaaaacacca cagaaactca gacctcaacc 660 tcaacttcca gccttcqttc agatcctaaa ctttqqttqq cqttqqqqac tqttqctaca 720 ggtctgatag ggttggcggc gacgggtatt gtacaggcgc ttgcattgac gccggagccg 780 gatageceaa ceaegacega ceetgatgea getgeaagtg caaetgaaac tgegacaaga 840 gatcagttaa cgaaagaagc gttccagaac ccagataatc aaaaagttaa tatcgatgag 900 ctcqqaaatq cqattccqtc aqqqqtattq aaaqatqatq ttqttqcqaa tataqaaqaq 960 caggetaaag cagcaggega agaggecaaa cagcaageca ttgaaaataa tgetcaggeg 1020 caaaaaaaat atgatgaaca acaagctaaa cgccaggagg agctgaaagt ttcatcgggg 1080 gctggctacg gtcttagtgg cgcattgatt cttggtgggg gaattggtgt tgccgtcacc 1140 gctgcgcttc atcgaaaaaa tcagccggta gaacaaacaa caacaactac tactacaact 1200 acaactacaa gcgcacgtac ggtagagaat aagcctgcaa ataatacacc tgcacagggc 1260 aatgtagata cccctgggtc agaagatacc atggagagca gacgtagctc gatggctagc 1320 acctegtega etttetttga eaetteeage atagggaeeg tgeagaatee gtatgetgat 1380 gttaaaacat cgctgcatga ttcgcaggtg ccgacttcta attctaatac gtctgttcag 1440 aatatgggga atacagattc tgttgtatat agcaccattc aacatcctcc ccgggatact 1500 actgataacg gcgcacggtt attaggaaat ccaagtgcgg ggattcaaag cacttatgcg 1560 cqtctqqcqc taagtqqtqq attacqccat qacatqqqaq qattaacqqq qqgqaqtaat 1620 agegetgtga atacttegaa taacceacca gegeegggat eecategttt egtetaaata 1680 tatccataat cattttattt agagggaggg aggggggaag tct 1723

<210> 4

<211> 559

<212> PRT

<213> Escherichia coli

<400> 4

Met Pro Ile Gly Asn Leu Gly His Asn Pro Asn Val Asn Asn Ser Ile
1 5 10 15

Pro Pro Ala Pro Pro Leu Pro Ser Gln Thr Asp Gly Ala Gly Gly Arg
20 25 30

Gly Gln Leu Ile Asn Ser Thr Gly Pro Leu Gly Ser Arg Ala Leu Phe 35 40 45

Thr Pro Val Arg Asn Ser Met Ala Asp Ser Gly Asp Asn Arg Ala Ser 50 55 60

Asp Val Pro Gly Leu Pro Val Asn Pro Met Arg Leu Ala Ala Ser Glu 65 70 75 80

Ile Thr Leu Asn Asp Gly Phe Glu Val Leu His Asp His Gly Pro Leu 85 90 95

Asp Thr Leu Asn Arg Gln Ile Gly Ser Ser Val Phe Arg Val Glu Thr

			100					100					110		
Gln	Glu	Asp 115	Gly	Lys	His	Ile	Ala 120	Val	Gly	Gln	Arg	Asn 125	Gly	Val	Glu
Thr	Ser 130	Val	Val	Leu	Ser	Asp 135	Gln	Glu	Tyr	Ala	Arg 140	Leu	Gln	Ser	Ile
Asp 145	Pro	Glu	Gly	Lys	Asp 150	Lys	Phe	Val	Phe	Thr 155	Gly	Gly	Arg	Gly	Gly 160
Ala	Gly	His	Ala	Met 165	Val	Thr	Val	Ala	Ser 170		Ile	Thr	Glu	Ala 175	Arg
Gln	Arg	Ile	Leu 180	Glu	Leu	Leu	Glu	Pro 185	Lys	Gly	Thr	Gly	Glu 190	Ser	Lys
Gly	Ala	Gly 195	Glu	Ser	Lys	Gly	Val 200	Gly	Glu	Leu	Arg	Glu 205	Ser	Asn	Ser
Gly	Ala 210	Glu	Asn	Thr	Thr	Glu 215	Thr	Gln	Thr	Ser	Thr 220	Ser	Thr	Ser	Ser
Leu 225	Arg	Ser	Asp	Pro	Lys 230	Leu	Trp	Leu	Ala	Leu 235	Gly	Thr	Val	Ala	Thr 240
Gly	Leu	Ile	Gly	Leu 245	Ala	Ala	Thr	Gly	Ile 250	Val	Gln	Ala	Leu	Ala 255	Leu
Thr	Pro	Glu	Pro 260	Asp	Ser	Pro	Thr	Thr 265	Thr	Asp	Pro	Asp	Ala 270	Ala	Ala
Ser	Ala	Thr 275	Glu	Thr	Ala		Arg 280	-	Gln	Leu		Lys 285	Glu	Ala	Phe
Gln	Asn 290	Pro	Asp	Asn	Gln	Lys 295	Val	Asn	Ile	Asp	Glu 300	Leu	Gly	Asn	Ala
Ile 305	Pro	Ser	Gly	Val	Leu 310	Lys	Asp	Asp	Val	Val 315	Ala	Asn	Ile	Glu	Glu 320
Gln	Ala	Lys	Ala	Ala 325	Gly	Glu	Glu	Ala	Lys 330	Gln	Gln	Ala	Ile	Glu 335	Asn
Asn	Ala	Gln	Ala 340	Gln	Lys	Lys	Tyr	Asp 345	Glu	Gln	Gln	Ala	Lys 350	Arg	Gln
Glu	Glu	Leu	Lys	Val	Ser	Ser	Gly	Ala	Gly	Tyr	Gly	Leu	Ser	Gly	Ala

355 360 365

Leu Ile Leu Gly Gly Gly Ile Gly Val Ala Val Thr Ala Ala Leu His 370 375 380

Arg Lys Asn Gln Pro Val Glu Gln Thr Thr Thr Thr Thr Thr Thr Thr Thr 385 390 395 400

Thr Thr Ser Ala Arg Thr Val Glu Asn Lys Pro Ala Asn Asn Thr
405 410 415

Pro Ala Gln Gly Asn Val Asp Thr Pro Gly Ser Glu Asp Thr Met Glu 420 425 430

Ser Arg Arg Ser Ser Met Ala Ser Thr Ser Ser Thr Phe Phe Asp Thr 435 440 445

Ser Ser Ile Gly Gly Pro Cys Arg Ile Arg Met Leu Met Leu Lys His 450 455 460

Arg Cys Met Ile Arg Arg Cys Arg Leu Leu Ile Leu Ile Arg Leu Phe 465 470 475 480

Arg Ile Trp Gly Ile Gln Ile Ser Val Val Tyr Ser Thr Ile Gln His
485 490 495

Pro Pro Arg Asp Thr Thr Asp Asn Gly Ala Arg Leu Leu Gly Asn Pro 500 505 510

Ser Ala Gly Ile Gln Ser Thr Tyr Ala Arg Leu Ala Leu Ser Gly Gly 515 520 525

Leu Arg His Asp Met Gly Gly Leu Thr Gly Gly Ser Asn Ser Ala Val $530 \hspace{1.5cm} 535 \hspace{1.5cm} 540$

Asn Thr Ser Asn Asn Pro Pro Ala Pro Gly Ser His Arg Phe Val 545 550 555

<210> 5

<211> 1460

<212> DNA

<213> Escherichia coli

<400> 5

aattetgttg etgatgetge tgattetegt geeagtgata tteeeggaet teetacaaat 60 ceaetgeget ttgetgegte egaggtatet ttgeatggtg egettgaagt tetteatgat 120 aaagggggge ttgataetet taaetetget attggatett egttatteeg tgttgaaact 180

cgggatgatg gcagccatgt tgctatcggg caaaaaaatg gcctcgagac cactgttgtt 240 ttaagtgagc aagagttttc tagcttacag tcccttgatc ctgaaggtaa aaacaaattt 300 qtatttactg qaggccqcgg tggcccaggg catgctatgg tcacggttgc ttcagatatc 360 gccgaagccc gtcagaggat aatagataaa ttagaaccaa aggatacaaa ggagacgaag 420 gagccagggg atccaaatag tggcgaggga aaaatcattg aaattcatac ctcaacctca 480 acttctagcc tccqtqcaqa tcctaaactt tqqttqtcat tqqqqactat tqctqcaqqt 540 ctgataggga tggctgcgac ggggattgca caggctgttg cgttgactcc agagccggat 600 gacccaatca ctaccgaccc tgatgctgca gcaaacacag ctgaagcagc ggcaaaagat 660 cagttaacga aagaagcatt ccagaaccca gataaccaga aagttaatat cgatgagaac 720 ggaaatgcaa ttccgtccgg ggaactaaaa qatgatgttg ttqcqcaaat agcaqaacaa 780 gctaaagcgg cgggtgaaca ggccagacag gaagctattg aaagtaattc tcaggcgcag 840 caaaaatatg atgaacagca tgctaaacgc gaacaggaaa tgtctctttc atcgggggtt 900 ggctacggta ttagtggtgc gctgattctt ggcgggggaa ttggtgccgg tgttactgct 960 gctcttcatc ggaaaaacca accggcagaa caaacaatca ctacacgtac ggtagtcgat 1020 aatcagccta cgaataacgc atctgcgcag ggcaatactg acacaagtgg gccagaagag 1080 tecceggega geagacgtaa ttegaatgee ageetegeat egaaegggte tgacacetee 1140 agcacgggca cggtagagaa tccgtatgct gacgttggaa tgcccagaaa tgattcactg 1200 gctcgcattt cagaggaacc tatttatgat gaggtcgctg cagatcctaa ttatagcgtc 1260 attcaacatt tttcagggaa cagcccagtt accggaaggt tagtgggaac cccagggcaa 1320 ggtatccaaa gtacttatgc gcttctggca agcagcggcg gattgcgttt aggtatggga 1380 ggattaacgg ggggtggcga gagcgcagta agtactgcca atgccqcacc aacgccqgqa 1440 cccgcacgtt tcgtttaaat 1460

<210> 6

<211> 484

<212> PRT

<213> Escherichia coli

<400> 6

Asn Ser Val Ala Asp Ala Ala Asp Ser Arg Ala Ser Asp Ile Pro Gly
1 5 10 15

Leu Pro Thr Asn Pro Leu Arg Phe Ala Ala Ser Glu Val Ser Leu His
20 25 30

Gly Ala Leu Glu Val Leu His Asp Lys Gly Gly Leu Asp Thr Leu Asn 35 40 45

Ser Ala Ile Gly Ser Ser Leu Phe Arg Val Glu Thr Arg Asp Asp Gly 50 55 60

Ser His Val Ala Ile Gly Gln Lys Asn Gly Leu Glu Thr Thr Val Val 65 70 75 80

Leu Ser Glu Gln Glu Phe Ser Ser Leu Gln Ser Leu Asp Pro Glu Gly
85 90 95

Lys Asn Lys Phe Val Phe Thr Gly Gly Arg Gly Gly Pro Gly His Ala Met Val Thr Val Ala Ser Asp Ile Ala Glu Ala Arg Gln Arg Ile Ile Asp Lys Leu Glu Pro Lys Asp Thr Lys Glu Thr Lys Glu Pro Gly Asp Pro Asn Ser Gly Glu Gly Lys Ile Ile Glu Ile His Thr Ser Thr Ser Thr Ser Ser Leu Arg Ala Asp Pro Lys Leu Trp Leu Ser Leu Gly Thr Ile Ala Ala Gly Leu Ile Gly Met Ala Ala Thr Gly Ile Ala Gln Ala Val Ala Leu Thr Pro Glu Pro Asp Asp Pro Ile Thr Thr Asp Pro Asp Ala Ala Ala Asn Thr Ala Glu Ala Ala Ala Lys Asp Gln Leu Thr Lys Glu Ala Phe Gln Asn Pro Asp Asn Gln Lys Val Asn Ile Asp Glu Asn Gly Asn Ala Ile Pro Ser Gly Glu Leu Lys Asp Asp Val Val Ala Gln Ile Ala Glu Gln Ala Lys Ala Ala Gly Glu Gln Ala Arg Gln Glu Ala Ile Glu Ser Asn Ser Gln Ala Gln Gln Lys Tyr Asp Glu Gln His Ala Lys Arg Glu Gln Glu Met Ser Leu Ser Ser Gly Val Gly Tyr Gly Ile Ser Gly Ala Leu Ile Leu Gly Gly Gly Ile Gly Ala Gly Val Thr Ala Ala Leu His Arg Lys Asn Gln Pro Ala Glu Gln Thr Ile Thr Thr Arg Thr Val Val Asp Asn Gln Pro Thr Asn Asn Ala Ser Ala Gln Gly Asn

Thr Asp Thr Ser Gly Pro Glu Glu Ser Pro Ala Ser Arg Arg Asn Ser 355 360 365

Asn Ala Ser Leu Ala Ser Asn Gly Ser Asp Thr Ser Ser Thr Gly Thr 370 380

Val Glu Asn Pro Tyr Ala Asp Val Gly Met Pro Arg Asn Asp Ser Leu 385 390 395 400

Ala Arg Ile Ser Glu Glu Pro Ile Tyr Asp Glu Val Ala Ala Asp Pro 405 410 415

Asn Tyr Ser Val Ile Gln His Phe Ser Gly Asn Ser Pro Val Thr Gly 420 425 430

Arg Leu Val Gly Thr Pro Gly Gln Gly Ile Gln Ser Thr Tyr Ala Leu 435 440 445

Leu Ala Ser Ser Gly Gly Leu Arg Leu Gly Met Gly Gly Leu Thr Gly 450 455 460

Gly Gly Glu Ser Ala Val Ser Thr Ala Asn Ala Ala Thr Pro Gly Pro 465 470 475 480

Ala Arg Phe Val

<210> 7

<211> 30

<212> PRT

<213> Escherichia coli

<400> 7

Pro Ile Gly Asn Leu Gly Asn Asn Val Asn Gly Asn His Leu Ile Pro 1 5 10 15

Pro Ala Pro Pro Leu Pro Ser Gln Thr Asp Gly Ala Ala Arg 20 25 30

<210> 8

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

Sequence

<400>	8	
aaagtc	gaca agaacctgag aaccag	26
<210>	9	
<211>	30	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Description of Artificial Sequence: Primer	
	Sequence .	
<400>	9	
tttata	gact tatgettgtg aaggtagtgg	30